Foundations for a Virtual Heliospheric Observatory:
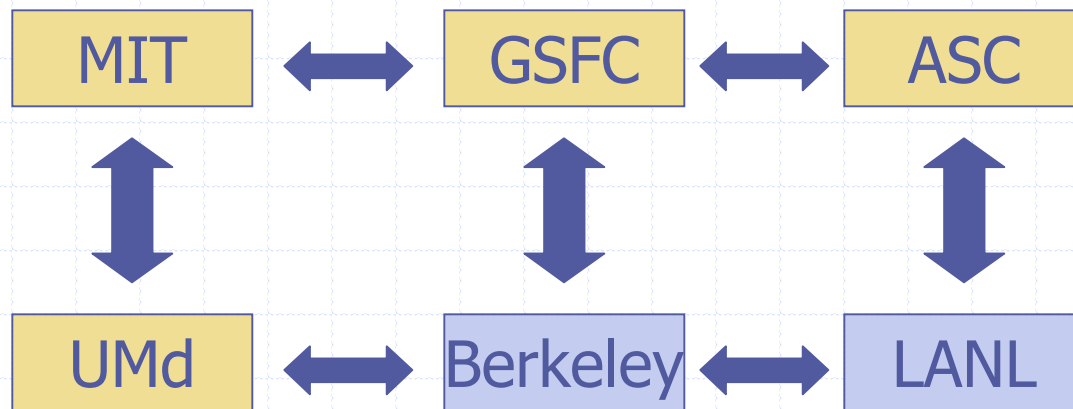
United L1 data sets, prototypical searching and distributed data processing

T.W. Narock[1,2], A. Szabo[2], A. Davis[3], R. Boller[2]
tom.narock@gsfc.nasa.gov

1. L3 Communications, GSI
2. NASA/Goddard Space Flight Center
3. ACE Science Center, Caltech

Foundations of the
Virtual Heliospheric Observatory (VHO)

Data Synchronization

Data is synchronized autonomously among all participating data providers using Rsync software

MIT ⟷ GSFC ⟷ ASC

⬍ ⬍ ⬍

UMd ⟷ Berkeley ⟷ LANL

GSFC – Goddard Space Flight Center – WIND MFI Data

MIT – Massachusetts Institute of Technology – WIND SWE Data

UMd – University of Maryland, College Park – SOHO Celias Data

Berkeley – University of California, Berkeley – WIND 3DP Data

ASC – ACE Science Center, California Institute of Technology – ACE MAG and SWEPAM Data

LANL – Los Alamos National Laboratory – Genesis Data

▢ Implemented
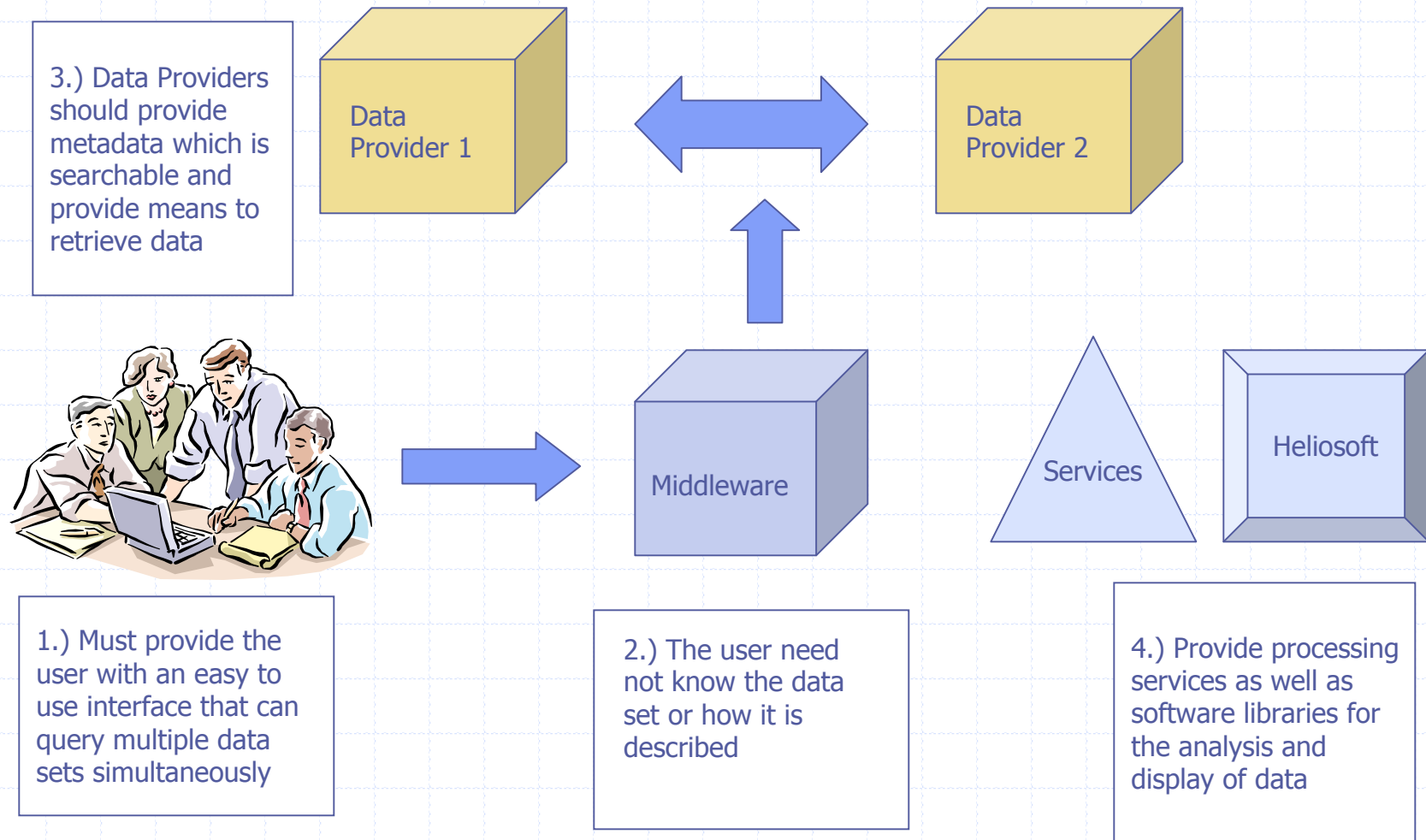
▢ In process of being implemented

Foundations of the
Virtual Heliospheric Observatory (VHO)

Data Synchronization (Continued)

Benefits to the Community

1.) Allows creation of new data products

- example: Merged WIND MFI and SWE data

http://lepmfi.gsfc.nasa.gov/plots.html

2.) Principle investigators can use other data sets

to calibrate their own instruments thus providing

higher quality data

VHO Prototype - Requirements

3.) Data Providers should provide metadata which is searchable and provide means to retrieve data

Data Provider 1

Data Provider 2

Middleware

Services

Heliosoft

1.) Must provide the user with an easy to use interface that can query multiple data sets simultaneously

2.) The user need not know the data set or how it is described

4.) Provide processing services as well as software libraries for the analysis and display of data
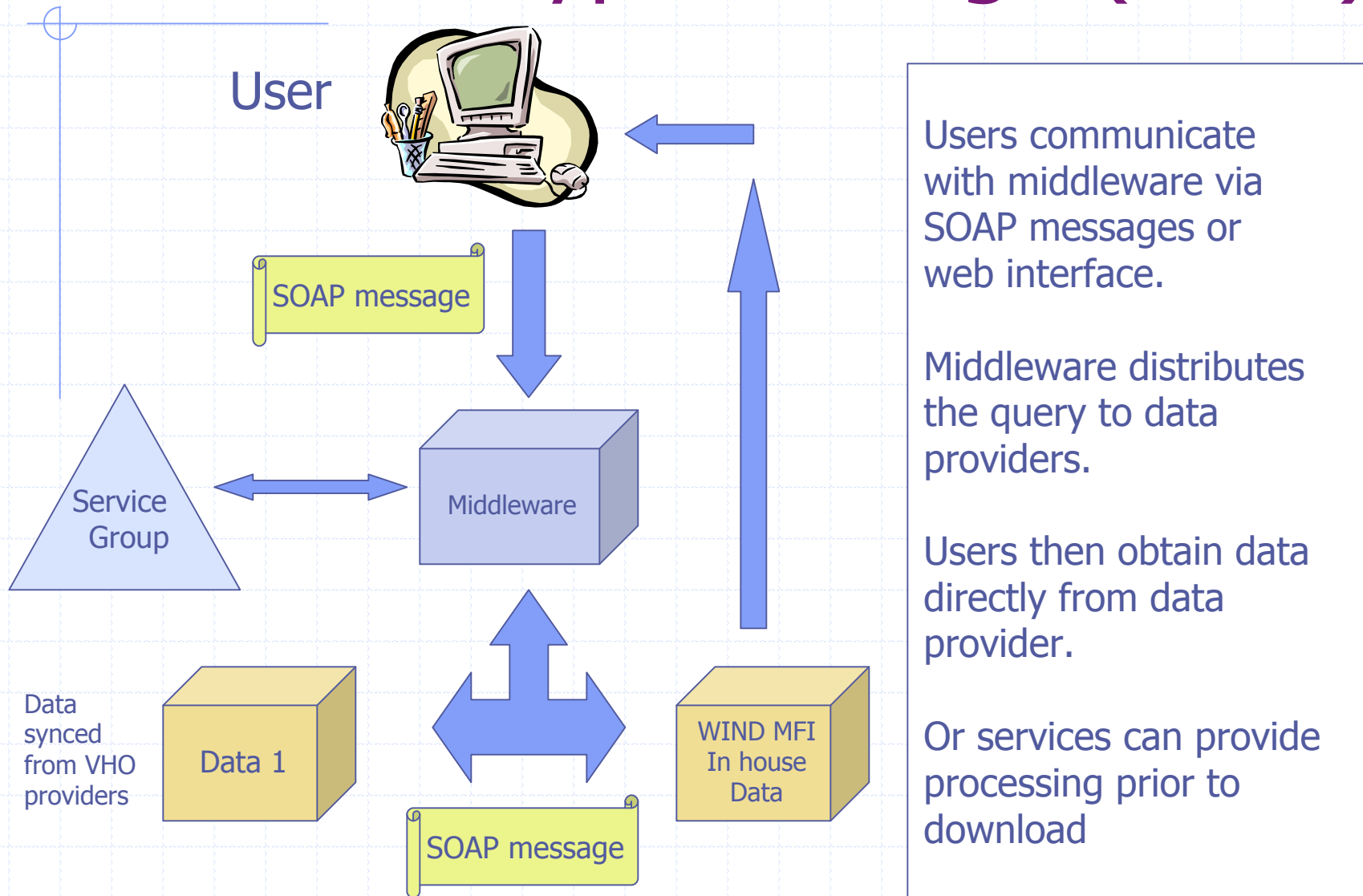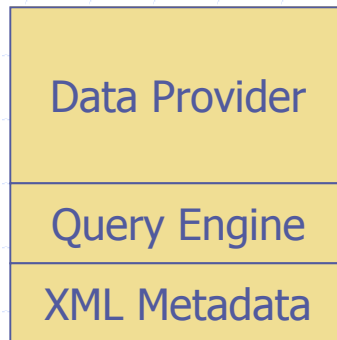
VHO Prototype - Design

A Prototype VHO was developed at GSFC

- Being tested at GSFC to later be deployed nationally

- It utilizes all datasets which are currently synced to GSFC and provides access to most recent data

- Provides one interface to many data sets, via web access or stand alone program

- Provides search and retrieval capabilities for Heliospheric data, allows searches by date, instrument and/or spacecraft position

- Lays the foundations for services and presently provides first service - coordinate transformations

VHO Prototype - Design (cont.)

User

SOAP message

Service Group

Middleware

Data synced from VHO providers

Data 1

SOAP message

WIND MFI In house Data

Users communicate with middleware via SOAP messages or web interface.

Middleware distributes the query to data providers.

Users then obtain data directly from data provider.

Or services can provide processing prior to download

Prototype Queries

| |
|---|
| Data Provider |
| Query Engine |
| XML Metadata |

- Query Engine sits on Data Provider and listens for incoming SOAP messages.

- Search Parameters are removed from message and metadata is searched

- XML metadata consists of one metadata file for each data file. Files are small and contain date and min/max values of spacecraft orbit over the time period covered by the corresponding data file

- An advanced search was implemented using higher level science terms such as searching for "L1 data", "solar wind data" or "inner heliosphere data"

Services

- Search and retrieval can only go so far

- Users need more than access to data

- Services will provide commonly used processing capabilities

- They will save the community time and effort and provide data in a form ready for use

- There are multiple ways of integrating services – 3 promising methods are being investigated for:

1. Ease of use and expansion potential

2. Efficiency

3. Ease of Integration into VHO and potentially other VOs

First Service – Coordinate Transformations
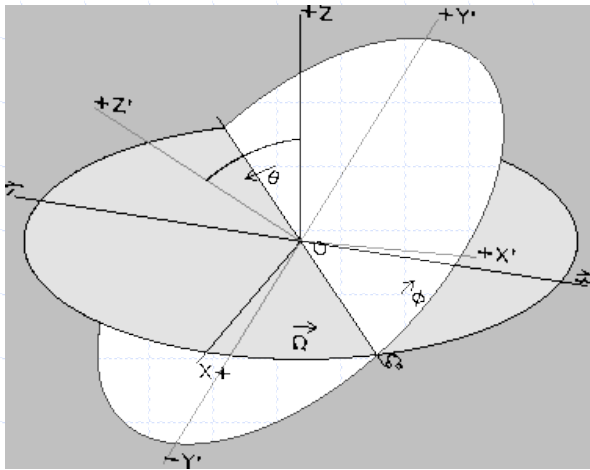
◈ Based on introductions and resources by Mike Hapgood, Markus Fraenz and Christopher Russel

http://sspg1.bnsc.rl.ac.uk/Share/Coordinates/ct_home.htm

http://www.space-plasma.qmul.ac.uk/heliocoords

http://www-ssc.igpp.ucla.edu/personnel/russell/papers/gct1.html

Testing against GEOPACK* shows results within +/- 0.001**

* Fortran transformation subroutines, implemented at http://sscweb.gsfc.nasa.gov

** When compared using same IGRF coefficients. At last check SSCWEB was using older coefficients than current version of this service

Services – Design 1

- Designed as a peer-to-peer network

- Services are more dynamic than data providers

- Multiple peers may offer similar services

- VxOs can join peer group and use services (VHO Middleware is a peer)

- Peer group is closed, i.e. must be approved user to contribute services

VHO
Middleware

Service
Peer 1

Service
Peer 2
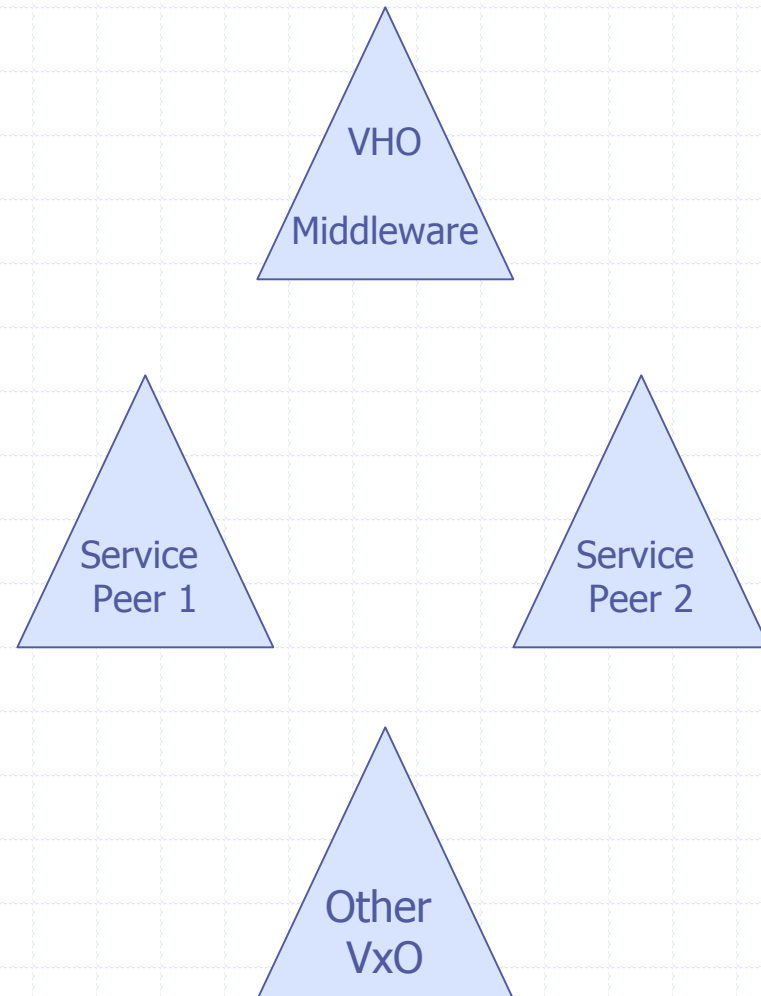
Other
VxO

Services – Implementation 1
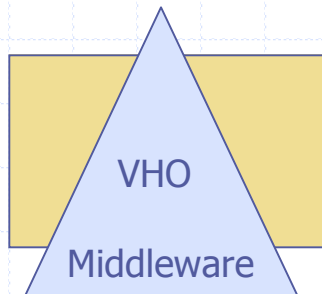
VHO Middleware

- VHO middleware understands SOAP messages and is also a peer in the service peer group

- Allows VHO middleware to quickly and easily find and use services

- Users need not learn another interface. To request a service they communicate with middleware using SOAP as they would if they were sending a data request (only inputs in SOAP message change slightly)

- Additionally, some common services will also have a web interface for easy use by non-programmers

Services – Technical 1

- Service Peer Group Implemented using Sun's JXTA Protocols

- JXTA allows peers to communicate directly and share resources even though some peers are behind firewalls or using different network transports

- Service Peers retrieve data on their own or by using VHO

- First service obtains data through direct access using APIs developed by ACE Science Center

Services – Layout 1

Service
Peer

Service
Peer

JXTA
Relay
Peer

Service
Peer

GSFC Firewall

Service
Peer

VHO
Middleware

GSFC Firewall
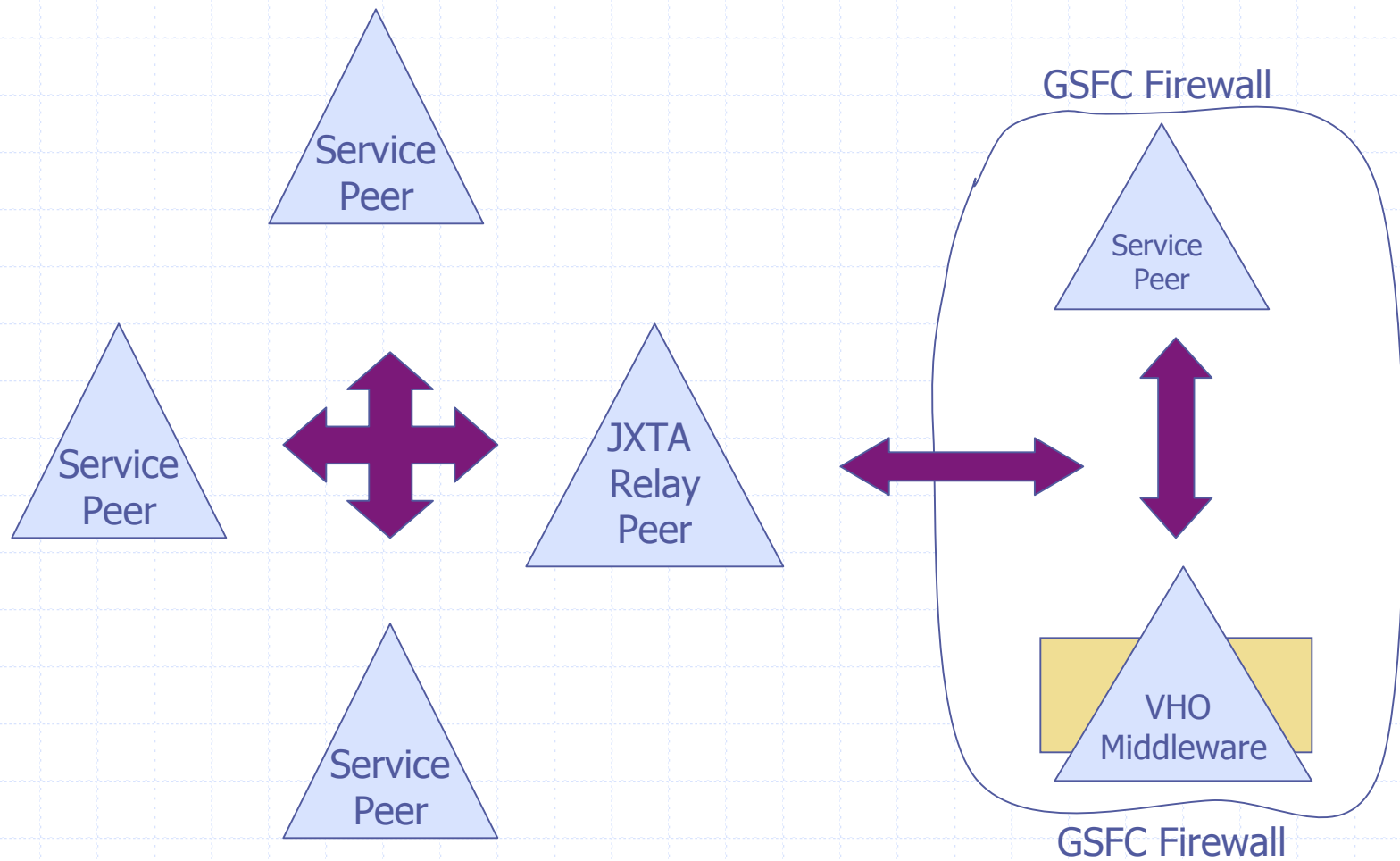
Why not have it all P2P?

- JXTA does not guarantee messages will arrive or be responded to

- Requires an all Java implementation and forces users to implement their own VxO software in Java with extensive JXTA

- New JXTA releases are frequent and some have been found to be worse in performance
[Seigneur et. al., ACM International Conference 2003]

- Unnecessary overhead (both programmatically and in network traffic) in querying data

- Data providers are static and simple message passing, which SOAP provides, is sufficient

Services – Design 2 and 3

- Create services that are compatible with the existing Collaborative Sun-Earth Connector (CoSEC) framework
- Provide stand alone application with plug-in services
- Or a possible mixture of both approaches

Services – Implementations 2
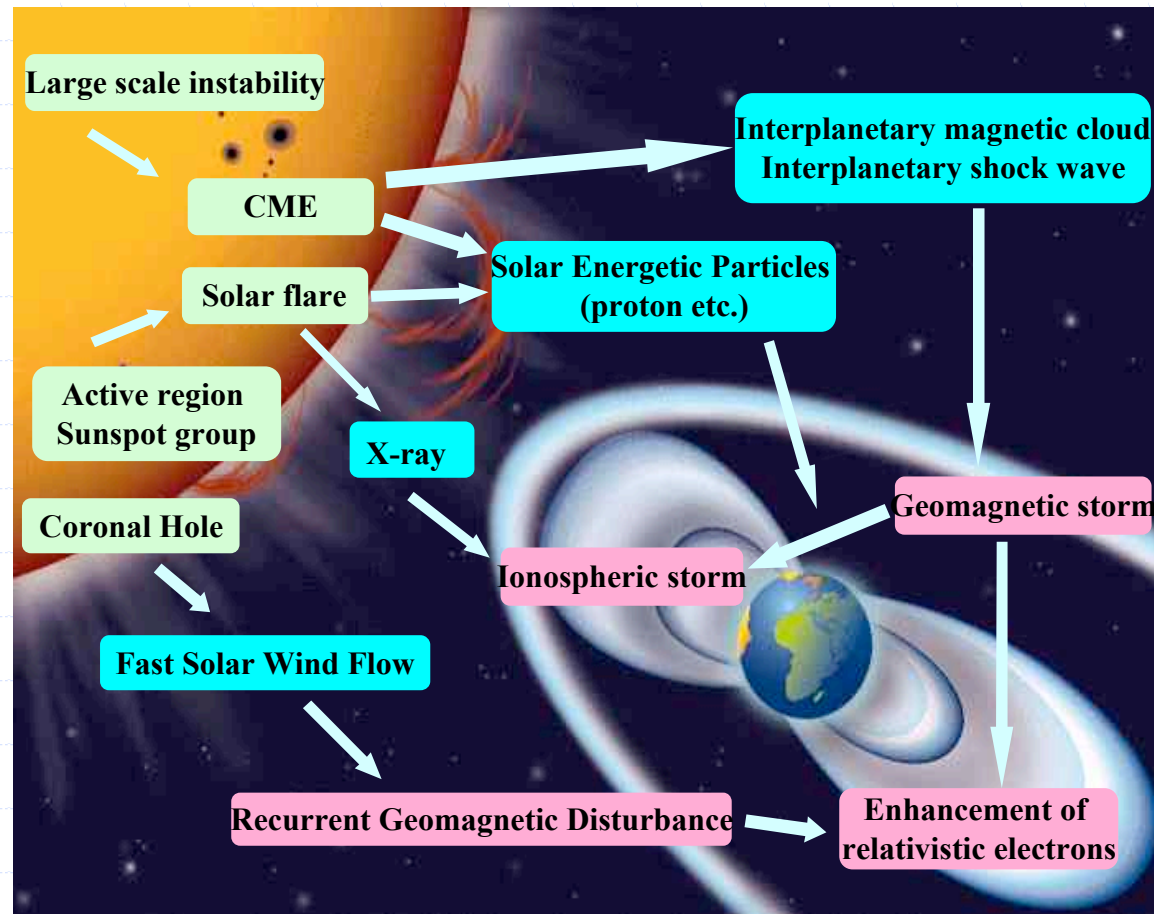
- CoSEC began as a solar initiative
- Explore integrating heliospheric data and services and allow cross-discipline science

Future Plans

- Add more Heliospheric data sets and services

- Finalize prototype and distribute to PI sites

- Finalize CoSEC service and explore collaboration

- Provide the community with a stand alone application that utilizes plug in services and possibly access to CoSEC
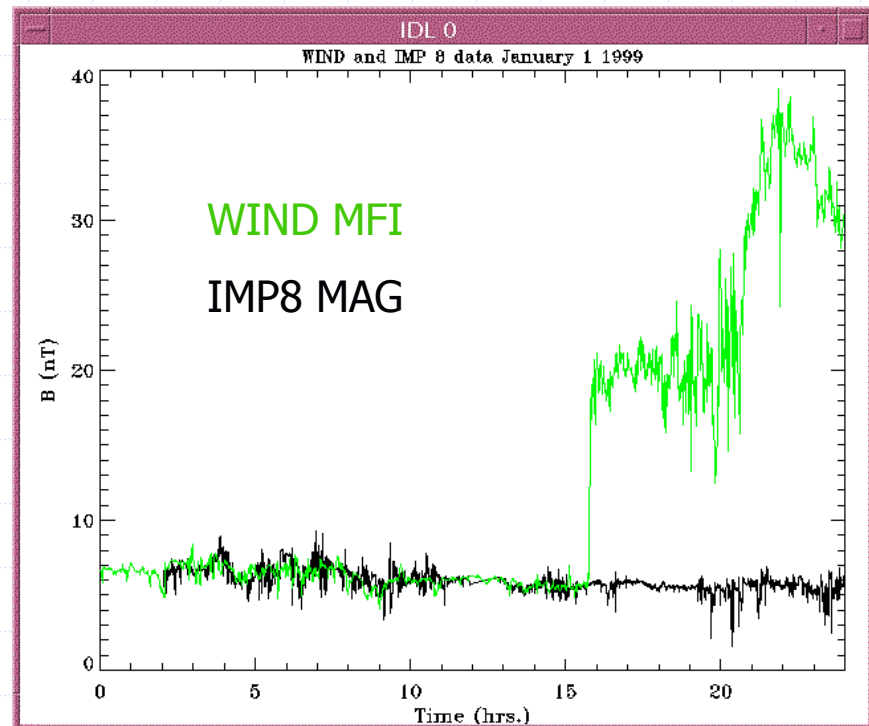
What is needed to participate?

Basic User

- Web Browser

Advanced User

- Programming experience

- SOAP library (available for most languages – C, Perl, Java, etc.

Data Provider

- SOAP daemon – provided by VHO team

- SOAP library – available free online

- Perl XML software – available free from CPAN

- Query Engine – provided by VHO team

Additional Information and Resources

◆ VHO home page – http://vho.nasa.gov

◆ VHO Prototype basic search page
http://vho.nasa.gov/search.html

◆ VHO API information
http://vho.nasa.gov/api.html

◆ VHO Advanced search and service documentation
http://vho.nasa.gov/services.html